

From Data to Discovery

HOW do you design a computer system that will be relevant and useful for decades? For computer architects developing the control system for Lawrence Livermore's National Ignition Facility (NIF), the central nervous system of the world's largest and most energetic laser, the solution centered on designing flexibility into the system from the beginning. The size of three football fields and 60 times more powerful than any other laser, NIF was declared fully operational in March 2009. It is run by the most complex, real-time control system ever designed for scientific research, a system that is intimately involved in all aspects of NIF's performance and maintenance.

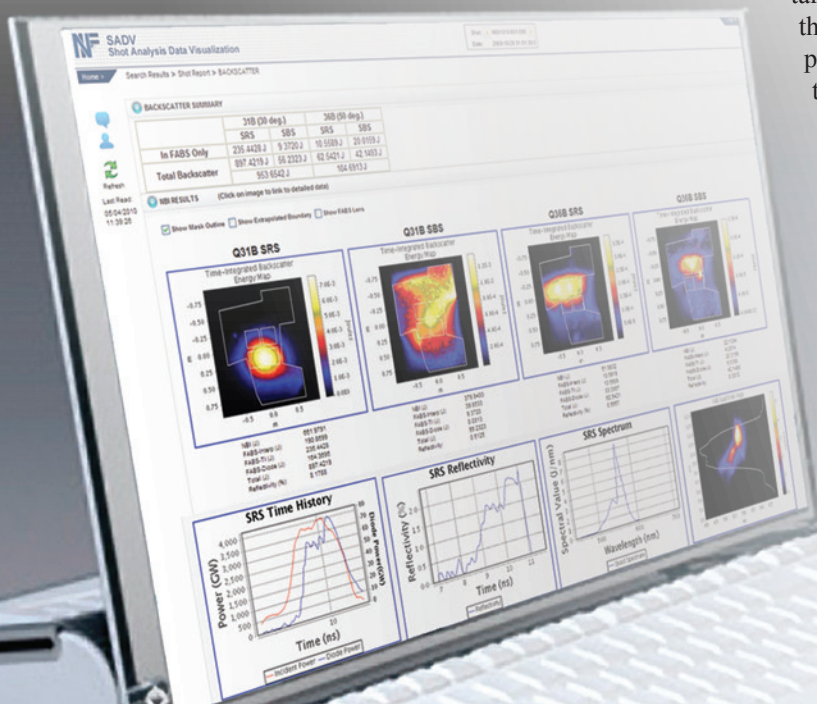
The behind-the-scenes workhorse, the integrated computer control system (ICCS), was designed and developed by NIF's Controls and Information Systems (CIS) Group. Paul Van Arsdall and Larry Lagin lead the team of about 200 developers, engineers,

technicians, information technology specialists, and quality control experts. According to Lagin, a 30-year fusion technology veteran, the key attribute of the control system is its dynamic, flexible design. Both the design of ICCS and the goals of the CIS Group have had to be flexible to meet the facility's evolving needs and to support its users. The original focus of control system software and hardware development was building functionality to enable NIF to successfully meet its milestones throughout the construction project.

Aiming for Automation and Flexibility

The controls effort has shifted over the last several years to developing software for automating experiments. That software, originally deployed to automate shots to commission the laser, has been expanded to control the full range of experiments planned for the National Ignition Campaign. Capabilities added to the control system over the past year include positioner control for aligning targets, cryogenic target systems for creating ice layers inside the targets, target diagnostics systems for determining reaction properties, and numerous systems for supporting the use of tritium and other target materials.

Now in its fourteenth year of software deployment, the CIS Group continues to upgrade control system software and add automation and new capabilities. Because the control system is modular and flexible, it can be upgraded piece by piece, without interrupting laser performance. Early on,





In the NIF Control Room, which is modeled after the National Aeronautics and Space Administration's mission control room, operations staff members interact with the "supervisory" portion of the control system. The sophisticated system is designed to use manual control and automation as applicable.

the CIS Group chose CORBA middleware as the architecture that communicates among various hardware and software platforms across the control system. According to Lagin, CORBA has proved to be flexible. "We can now upgrade languages and our computers because we chose CORBA," he says. Ada, the primary software language within the processors that interface with motors and other laser and target equipment, is no longer a common language choice. This change has prompted CIS engineers to gradually migrate to Java, a current industry-standard language that offers excellent diagnostic and development tools and can run on many types of operating systems. Because CORBA allows disparate languages to talk to one another, having some components in Ada and others in Java during the gradual migration is easily accommodated. The software language migration is just one aspect of the ongoing control system software and hardware maintenance effort that happens behind the scenes.

As NIF moves toward full operation as a scientific user facility with an anticipated 30-year lifespan, the CIS Group also is endeavoring to increase the number of tools that will support science experiments and meet the needs of current and future NIF users. Ric Beeler leads a team that has been expanding and improving these tools. Scientists performing an experiment at NIF first interact with campaign management tools, a suite of

software applications designed to translate experimental plans and specifications into actions for the control system. After a shot, scientists use the shot analysis, visualization, and infrastructure (SAVI) tools to archive, analyze, and view the experimental results. While a version of the campaign management tools has existed since early experiments were performed in 2003 with just four of NIF's 192 beams, the system is much more sophisticated now. "The management tools have grown up with the control system," says Beeler. Several tools have recently been added to the suite to help scientists more effectively set up their desired shots.

Mining Experimental Data for Gems

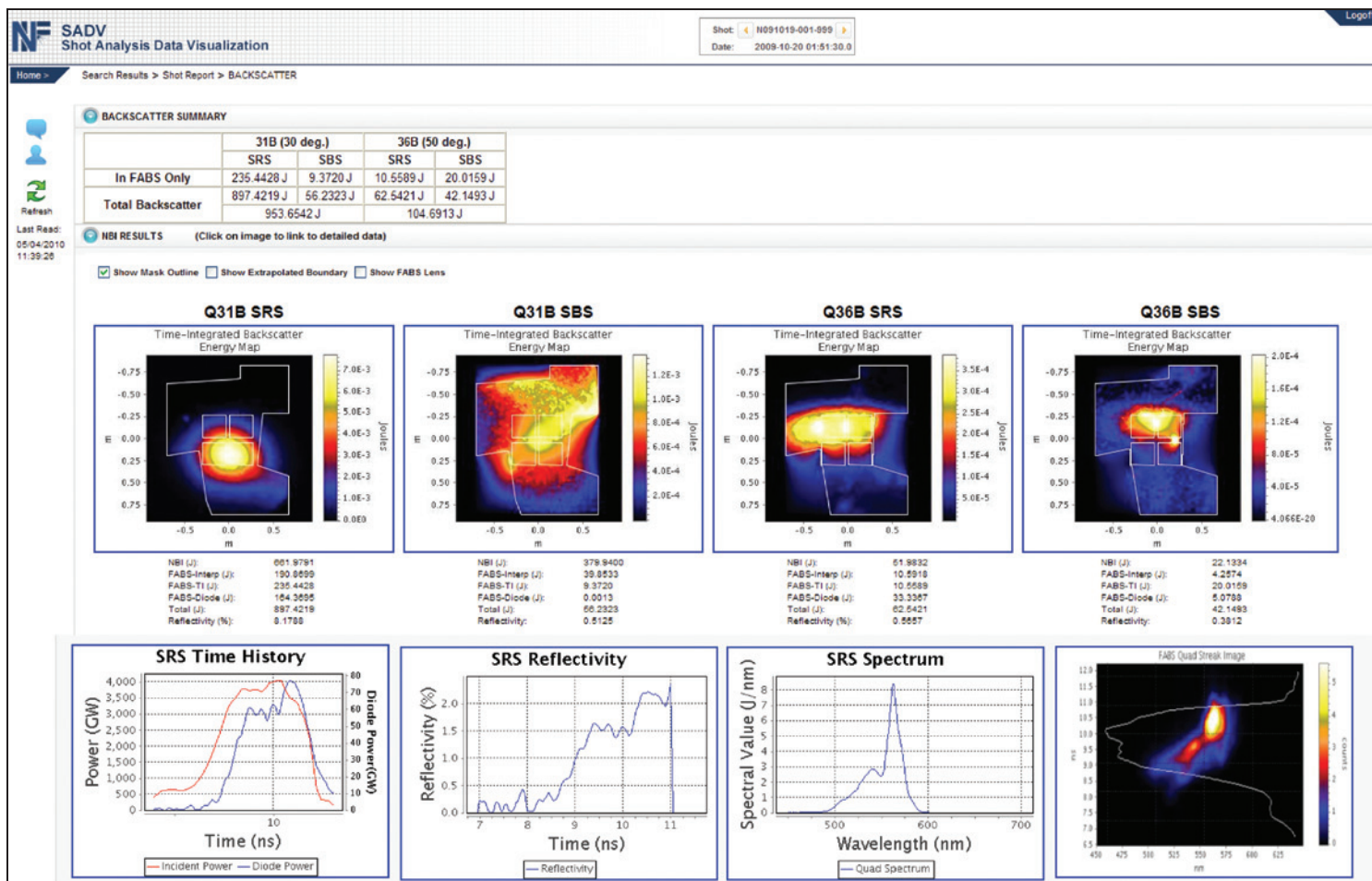
The speed and power of NIF's 192 laser beams converging on and bombarding a tiny target with massive energy may inspire wonder, but what happens after a NIF experiment is no less crucial to advancing science and energy research. Managing, organizing, and storing the flow of data from these experiments is no small task. Ten terabytes of experimental data, in multiple formats and from various sources, are generated each month, an amount equivalent to the entire printed collection of the U.S. Library of Congress. Data generation actually begins before a laser shot. Information is captured about the initial condition of each experiment, including the precise shape of the target capsule and

the thickness of ice layers within a cryogenic target. (See *S&TR*, June 2010, pp. 17–19.) Immediately after a laser shot, SAVI automatically records, processes, and refines much of the raw data produced by the dozens of x-ray, nuclear, and optical diagnostics that measure aspects of the experiment. (See *S&TR*, December 2010, pp. 12–18.)

The SAVI tools were developed in close collaboration with scientists who perform target experiments to determine how the software tools could best support NIF users. SAVI archives all of the target diagnostic data that are captured electronically by ICCS. A portion of the archived data is processed through the SAVI analysis pipeline, while the remaining data are made available to scientists for use in their analysis applications. The SAVI analysis engine processes information at three hierarchical levels representing the steps in the analysis flow: instrument, diagnostic, and campaign. For example, at the instrument level, SAVI will

provide results from an individual hardware device, such as a camera, adjusting for instrument performance factors such as nonlinear gains. At the campaign level, it will refine and combine data from several diagnostics to determine overall characteristics of the experiment such as radiation temperature.

SAVI data visualization tools help scientists view and understand the experimental results. Within these Web applications, scientists can choose to view the raw data or any level of the analysis. Researchers want to access experimental data soon after a laser shot, with the minimum likelihood of error. They use some of these data to determine parameters for the next shot in a series, which might occur just a few hours later. SAVI tools provide experimental results within 30 minutes of a shot through a Web interface, with error bounds and quality metrics. Scientists can review the data remotely or locally, download results, and perform and upload their own analysis.



The shot analysis, visualization, and infrastructure tools help scientists quickly view and interpret data from a just-completed NIF experiment.



Tim Frazier inspects control system servers that run applications for recording and processing shot data. Frazier led the development of the NIF experimental data repository, which captures and stores the data produced by each laser shot.

Safeguarding Results for Years to Come

Experimental data, plus data on the postshot state of the facility, are housed in and retrieved from the NIF data repository. This archive, designed over the past few years, stores all the relevant experiment information—including target images, diagnostic data, and facility equipment inspections—for 30 years using a combination of high-performance databases and archival tapes. Although most scientists analyze data and publish results shortly after an experiment, retaining the data will allow researchers to retroactively analyze and interpret results as scientific fields advance and theories change, or perhaps to build on experimental data originally produced by other scientists.

The intention is to eventually make much of the unclassified data available to the scientific community at large. Beeler notes that this wealth of stored experimental data will make the archive “a real treasure.” The staff maintaining the archive faces the ongoing challenge of keeping the archive’s hardware and software modern, functional, and relevant. “It’s a constant race to keep up to date but still maintain a stable system,” says Tim Frazier, head of NIF Information Technology and the data architect who led the development of NIF data storage capabilities.

A crucial design feature of the database, from a scientist’s point of view, is preserving the pedigree of the data. Frazier likens the data pedigree to a family tree that traces members and relationships through time. With science data, the “family members” are all the linked pieces of information from a particular experiment such as algorithms, equipment calibrations, configurations, images, and raw and processed data. If it were discovered months after an experiment that a camera was miscalibrated, for example, the correction could be entered into the data set, and all of the family members, or linked pieces of data that feed from or rely on that calculation, would be overlain by new, corrected versions. Having a long-term record of this linked, versioned data is invaluable to scientists as proof of the validity of their results.

The database relies on both commercial Oracle software and open-source technology. What sets it apart from other large scientific data archives, such as that of CERN (the European Organization for Nuclear Research), is that both structured and unstructured data reside inside the database. By bringing images and other raw, unstructured content into the database, CIS staff can use one tool set and one skilled team to back up and manage the data. According to Frazier, being able to handle these “blobs” of data is a fortunate matter of timing; the database was designed and built at a point where computer technology was fast and powerful enough to handle both raw and processed data files within one archive.

The control and information systems consist of more than 2,000 computers, terabytes of data, and plentiful software tools, all working together to produce, record, and process experimental information. As Frazier emphasizes, NIF is a marvel, but its purpose is simple. “NIF is an incredible data-generating machine,” he says. “What really matters are the data and the scientific understanding that evolves from this information.” Handling, processing, and storing experimental data for scientists to use and mine for new discoveries is a crucial charge for CIS staff, one that enables and supports the work of experimentalists. From these vast records made up of ones and zeroes will likely come groundbreaking discoveries and evidence to advance fusion research and many other areas of science.

—Rose Hansen

Key Words: campaign management tools; CORBA; computer automation; data archive; integrated computer control system (ICCS); National Ignition Facility (NIF); shot analysis, visualization, and infrastructure (SAVI).

For further information contact Trish Damkroger (925) 422-5816 (damkroger1@llnl.gov).